

観光ルート推薦アルゴリズムの検討

Examination of tour recommendation algorithm

金川 元紀
Geki Kanagawa

岡山大学 太田研究室
Ohta Laboratory, Okayama University

概要 観光客が観光を行う際、複数の観光スポットを訪問したい場合がある。中川らは訪れたい観光スポットを1つだけ指定可能な観光ルート推薦アルゴリズムを提案した。本稿では、中川らの観光ルート推薦アルゴリズムを改良し、複数の訪れたい観光スポットを指定可能とする拡張を行う。

1 はじめに

近年、東京や大阪など大都市への人口集中に伴い、地方の衰退が問題となっている。そこで、本研究では地方の観光を便利にする観光ルート推薦に注目する。中川ら[1]はツイートとヤフー知恵袋の情報を利用して適切な観光ルートを紹介するアルゴリズムを提案した。このアルゴリズムでは、出発地・出発時刻および到着地・到着時刻と、訪れたい観光スポット一つを入力としていた。しかし観光客が観光を行う際、複数の観光スポットを訪れたい場合も考えられる。本稿では、中川らの観光ルート推薦アルゴリズムに、訪れたい観光スポットを複数選択可能にする改良をする。

2 先行研究

2.1 観光スポットのスコア

新井ら[2]は観光スポットにスコアを与え、その合計が最大になるような観光ルートを紹介するアルゴリズムを提案した。中川ら[1]は新井らの提案したスコアにさらに、類似度スコア、NNカテゴリスコアの2つを追加した。新井らと中川らのアルゴリズムで利用されているスコアは以下のようである。

- 時間帯スコア：その観光スポットを訪れるのに良い時間帯を表すスコア
- 共起スコア：共に訪れられやすい観光スポットを表すスコア
- カテゴリスコア：旅行者がどういう目的で観光スポットを訪れているかを表すスコア
- 類似度スコア：観光スポットの特徴がどれ程似ているかを表すスコア
- NN カテゴリスコア：ニューラルネットワークにより、各観光スポットを「食事」、「景観」、「土産」、「行動」の4カテゴリに分類したそのスコア

このスコアの合計が最大になるような観光ルートを求めるのが新井らと中川らの観光ルート推薦アルゴリズムである。次に、スコアの合計が最大になるような観光ルートを求めるアルゴリズムの詳細を説明する。

2.2 ルート生成アルゴリズム

まず、被推薦者は訪れたい観光スポットと観光の出発地

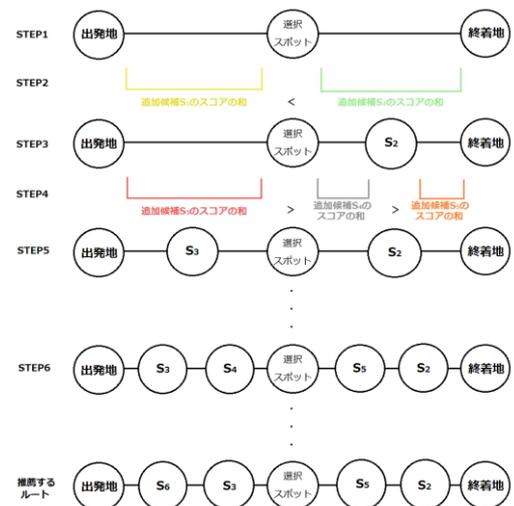


図1: ルート生成の流れ

・ 出発時刻および到着地・到着時刻を入力する。追加入力する観光スポットを選択スポットと呼ぶこととする。入力情報が確定した後に、次の計算で最適なルートを計算する。ここで図1に各計算ステップでのルートを示す。

- STEP1: 出発地→選択スポット→到着地の初期ルートを作成する。
- STEP2: 出発地→選択スポット, 選択スポット→到着地の各部分経路の間に訪れることができる観光スポットの中から、2.1節で説明したスコアの和が最大となる観光スポットを探す。図1ではS1とS2が各部分経路でスコアの和が最大の観光スポットである。
- STEP3: S1とS2でスコアの和が大きい方をルートに加え、出発地→選択スポット→S2→到着地のルートができる。
- STEP4: 再び、各部分経路で訪れることができるスポットの中からスコアの和が最大となるスポットを探す。
- STEP5: スコアの和が最大のスポットをルートに加える。
- STEP6: 加えられる観光スポットがなくなるまでSTEP4とSTEP5を繰り返してルートを生成した後、そのルートを保存する。

以上の計算だけでは、局所最適解に陥る可能性がある。そこで中川らは、ルートから観光スポットを無作為に削除し、STEP4とSTEP5を繰り返すことで新たなルートを生成し保存するという計算を追加した。そして互いに異なる100ルートが保存されるまで生成を繰り返した後、スコアの合計が最も高いルートを推薦ルートとした。

3 ルート生成アルゴリズムの変更

中川らのルート生成アルゴリズムの実装では、2.2 節の STEP1 において、入力スポットの情報を単純に文字列で表現していた。本研究では、入力スポットの情報をリストで表現する拡張を行った。この拡張により複数の選択スポットが含まれた初期ルートを生成できるような変更が可能となる。選択スポットの数が可変となったため、STEP2 や STEP4 でのスコア算出において、選択スポットと追加候補スポットにより算出される共起スコアと類似度スコアも同時に計算するよう処理を変更した。

4 評価実験

4.1 実験内容

訪れたい観光スポットを複数選択して観光ルート推薦を行う。その際、本当に良いルート推薦を行えているかを評価するために、出力された 100 個のルートのうちスコアの高いルートの上位 3 件を図に示し考察する。また、推薦ルートを出力するまでの時間に変化が出たかを調べるため、修正前のプログラムを用いて選択スポットを 1 つのみ選んだ時と、修正後のプログラムを用いて選択スポットを 3 つ選んだ時でルート推薦をそれぞれ 10 回ずつ行い、その計算にかかった時間の平均をとり比較する。

4.2 実験結果

出力されたルートの内、スコアが高いものの上位から図 2、図 3、図 4 に示す。3 位のルートは移動効率の良い遠回りのあるルートとなっている。1 位のルートについては、移動効率が良いルートとなっているが、イオンモール岡山と三井アウトレットパーク倉敷は同じ種類のショッピングスポットであるため同日に回る必要はないスポットである。そのため 2 位の推薦ルートの方が 1 位の推薦ルートよりも妥当なルートであると言える。

次に計算時間について評価する。選択スポットを 1 つ選んだ時の平均の計算時間が 699.4 秒、選択スポットを 3 つ選んだ時の平均の計算時間が 536.4 秒であった。選択スポットを 3 つ選んだ場合の方が平均の計算時間がやや少なくなる結果となった。

5 まとめ

本稿では、中川らのルート推薦アルゴリズムを改良し、複数の選択スポットを選択できるようにする改良を行った。しかし推薦ルートに同時に訪問するのが適切でないスポットが現れてしまう場合がある問題があることを確認した。この問題を解決するには、スコアについて検討し直す必要があると考えられる。計算時間については、複数の選択スポットを選択した場合に計算時間が短くなる傾向が見られた。これは複数のスポットを選ぶことでルートがある程度定められ、スポットを加えたり削除したりするループ処理の回数が少なくなるためである。



図 2：1 位の推薦ルート

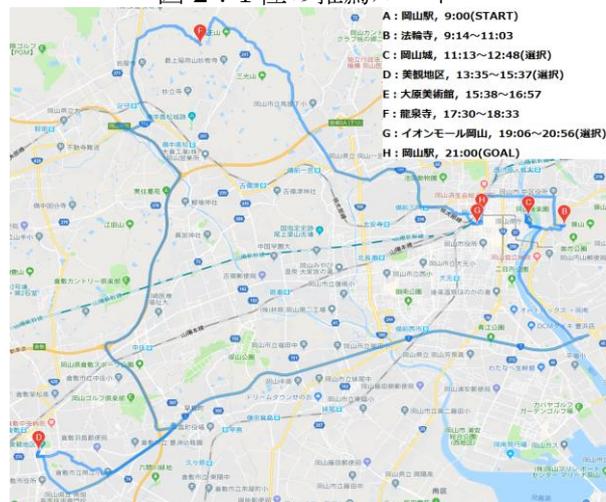


図 3：2 位の推薦ルート

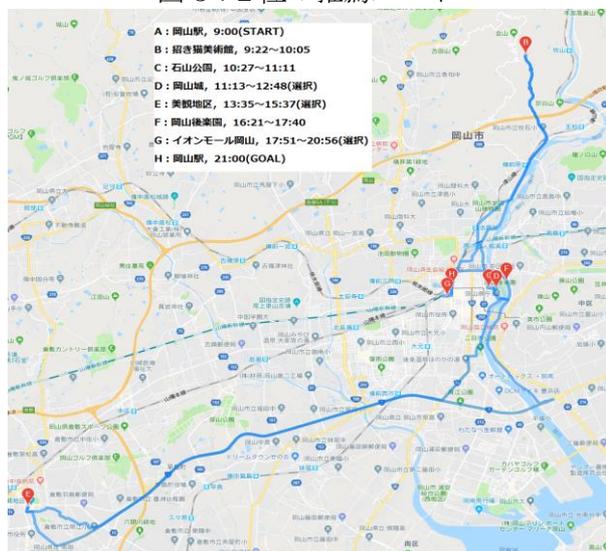


図 4：3 位の推薦ルート

参考文献

- [1] 中川智也, 新妻弘崇, 太田学, “分散表現を用いたパーソナライズド観光ルート推薦”, DEIM Forum 2018, H1-3, pp. 1-8, 2018.
- [2] 新井晃平, 新妻弘崇, 太田学, “Twitter を利用した観光ルート推薦の一手法”, DEIM Forum 2015, G7-6, pp. 1-8, 2015.